



# ON COPYRIGHT

## Copyright in the networked world: plagiarism and its ambiguities

Plagiarism and  
its ambiguities

691

Michael Seadle

*Humboldt University, Berlin, Germany*

Received 7 September 2008  
Accepted 7 September 2008

### Abstract

**Purpose** – Libraries and scholars face more frequent problems with and decisions about plagiarism than in the past. This article aims to look at complex cases where plagiarism may have occurred.

**Design/methodology/approach** – The method is anthropological and looks at specific cases, in which the situations are real but the actors have been fictionalised to protect identities.

**Findings** – Plagiarism tools, while invaluable for discovering potential problems, can also expose cases where judgments depend on complex circumstances.

**Originality/value** – The goal is to show areas where ambiguity in plagiarism cases exists.

**Keywords** Copyright law, Software tools

**Paper type** Case study

### Introduction

The digital world suffers from a plague of plagiarism. The non-digital past did too. The difference is not merely that copying is easier in the digital world, but that discovering plagiarism is also easier, thanks in part to tools like Turnitin.com[1], iThenticate.com[2] and Google (including Google Scholar). This means that libraries and scholars face more frequent decisions about plagiarism. This article looks at some of the more complex problems in recognising plagiarism. The simple issue of a student copying another person's whole paper and turning it in as their own is only the crudest form.

This article's methodology, as usual for me, is fundamentally anthropological, looking at both the law and institutional regulations in the context of social interactions. Because of the sensitivity of plagiarism accusations, any direct discussion of actual cases is clearly impossible. The individuals in the cases below are fictional, though the general situations have a basis in actual fact. The narratives are designed to represent problems, not people.

### Plagiarism and copyright

While many definitions of plagiarism exist, most agree fundamentally with the following version from the US Department of Health and Human Services' Office of Research Integrity (ORI):

As a general working definition, ORI considers plagiarism to include both the theft or misappropriation of intellectual property and the substantial unattributed textual copying of another's work (ORI, 1994).

Universities often have similar definitions in their codes. Humboldt-Universität zu Berlin (my own institution) defines plagiarism in its "Satzung über die Grundsätze der



Library Hi Tech  
Vol. 26 No. 4, 2009  
pp. 691-695

© Emerald Group Publishing Limited  
0737-8831  
DOI 10.1108/07378830810921049

Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis...” (“Charter on the Principles of Humboldt-Universität zu Berlin for the Assurance of good scholarly practice...” ) (Humboldt-Universität zu Berlin, 2000):

*§ 9 Verletzung geistigen Eigentums*

*Zur Verletzung geistigen Eigentums gehören insbesondere:*

*... die unbefugte Verwertung unter Anmaßung der Autorschaft (Plagiat) (p. 4).*

§9 Breach of intellectual property

To breaches of intellectual property belong especially:

... unauthorised use under the pretence of authorship (plagiarism) [my translation].

Plagiarism is not always a copyright violation. Plagiarism can occur when a work no longer has legal protection. In the USA that would be works published before 1923 and in the European Union (EU) it would be works by authors who had died more than 70 years earlier. But the act of copying without attribution is no less plagiarism when no copyright infringement is involved. It is certainly a violation of the rightful author’s “moral rights” by failing to recognise the right of attribution under EU laws. In the US, however, the right of attribution is limited to works of visual art[3].

The connection between plagiarism and copyright can mean that the rights holder has legal recourse via copyright law, but for most students and scholars, the consequences of plagiarism are harsher: expulsion for a student and loss of a job for others. The precise penalties vary. Often a jury style committee of peers or an institutional review board decide the actual punishment, but expulsion or firing inevitably rank among the real possibilities. Copyright infringement suits and settlements represent a separate and additional threat.

The cases below were chosen for their ambiguity.

### **Case 1: self-copying**

A successful and ethically conscientious author wrote to his editor about a problem when writing three simultaneous articles about different aspects of a recently completed research project. While the literature reviews and conclusions varied for each article, depending on the emphasis, the section describing the project’s research methodology was essentially the same in each article and almost had to be, because the author wanted to provide a consistent and accurate explanation of what went on in the project itself.

Since the articles were being sent to three different journals essentially simultaneously, the author could not cite a previously published version and for timeliness reasons did not want to wait for one publication to appear before sending out the others.

The editor agreed that repeating the methodology section word for word made good sense in this situation.

### **Case 2: self-plagiarism**

In this case an Asian author submitted an article to a journal that regularly uses iThenticate.com to check on potential plagiarism problems. With iThenticate the journal editor can set a copying threshold that sends an alert, which this article triggered. iThenticate also provides a comparison between the suspicious text and the text in the alternative source. In this case many of the suspicious passages came from other articles that the author had published.

English was not this author's native language. To some extent it was clear that he had borrowed sections from previously published works to save the trouble of having to reformulate parts of a literature review or methodology description in correct English. While this seems similar to the situation in case 1, it was not identical. This author's self-citation was not just in single section describing a factual set of research actions that ought to be described consistently, but included sections on results and conclusions where he had changed only key words about the topic of the paper.

iThenticate also matches quoted texts to their original source. This can be interesting for an editor who wants to see whether an author is taking a quotation out of context. In this case the editor noticed that key words had been changed in cited quotations by other authors. With some further research it became clear that the author had probably used a change-all command to alter a previous article to fit the new topic, and the change-all command changed all quoted texts in addition to the author's own.

It was impossible for the editor to tell whether the author deliberately falsified the quotations, or did it accidentally as part of copying and transforming his own text to the new topic. In this case the plagiarism tracking led to the equally serious problem of citation falsification. The author's self-citation raised some concern about language and the precision of the conclusions, but did not necessarily seem to require an immediate reject. The falsification of quotes by others was, however, too serious for the article not to be rejected.

### Case 3: sloppiness

In this case a US university librarian was asked to provide text for a short description of a foreign-language research archive for publication in a work that would go to about 400 people. Since she was not a particularly good or confident writer, she simplified the task by copying a two-paragraph English-language summary. The text was, however, too short, so she found additional information on the archive web site and translated it into English, in part with the help of online translation tools. This translated text was then coupled with the first and sent to the publisher with a note that included both original sources in the "references" section but with no quotation marks or indication about the translation.

The publisher recognised the source of the two-paragraph quote and accused the librarian very pointedly of plagiarism. She responded with a new copy that included the quotation marks. The publisher considered this inadequate and contacted the university, which by university policy triggered a formal hearing with a hearing board of five persons, including faculty and student representatives.

Apparently it genuinely never occurred to the librarian that the copying might be looked on as plagiarism. The forgotten quotation marks seemed to her like a trivial formatting detail, especially since the sources were in fact listed in the references. She had also not thought about the copyright infringement implicit in sending a publisher the two sections of text. She was more accustomed to the indirect process of sending materials to faculty who had requested them for their own articles than in preparing and publishing her own works.

One argument in favour of the librarian was that her sources did appear in the references and only the quotation marks were missing. The hearing board did not accept this argument, partly on the principle that the librarian had a Master's degree and ought to know better. Another argument was that the librarian had no intention of

benefiting from the plagiarism. Her name would appear on the publication only as someone who assisted in putting it together. The problem was that it was a breach of trust that could embarrass the publisher if discovered later. The hearing board decided that plagiarism had occurred, but let the librarian off with a warning because they believed ultimately that sloppiness was the cause, not an intention to plagiarise.

---

**Case 4: common discourse**

This case involves a well-published European author and an article on a widely discussed topic. iThenticate triggered an alarm because of a large percentage of text that matched text published elsewhere. In some places the text came from the author's own works. In other places the problem came from long phrases that had become standard jargon within the author's field. And in some cases the text, while essentially identical, carried little intellectual weight and seemed like standardised phrases that, within the context of that field, anyone might use.

No single item of text would trigger any plagiarism concern, but the sheer quantity of common or (potentially) copied phrases made the journal editor wonder to what degree the author's articles represented a simple cut and paste of established ideas. While it looked as if the author were potentially copying heavily from others in the field, it occurred to the editor that the same might be true for every author in the field because of the common language of discourse. He decided to accept the paper as the author's own work.

**Conclusion**

One goal of this article is to alert both authors and editors of the potential problems that arise from the use of contemporary plagiarism discovery tools. Authors like those in cases 2 and 3 need to be more careful and to think more like the author in case 1, who addressed potential problems intelligently and proactively. Editors and others need also to remember the danger of false positives in plagiarism searches, such as in case 4.

These four cases by no means represent the full spectrum of plagiarism problems and the intent is not to suggest particular ethical standards. Anthropological methodology is descriptive, not prescriptive. The intent instead is to show areas where ambiguity exists in how authors and institutions address plagiarism issues.

**Notes**

1. <http://turnitin.com/static/home.html>
2. <http://www.ithenticate.com>
3. 17 USC 107, *United States Code*, Title 17, Chapter 1, section 106a. Available at: [www.copyright.gov/title17/92chap1.html#106a](http://www.copyright.gov/title17/92chap1.html#106a) (September 2008).

**References**

- Humboldt-Universität zu Berlin (2000), "Satzung über die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis und über den Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens", available at: [www.amb.hu-berlin.de/2002/33/3320020](http://www.amb.hu-berlin.de/2002/33/3320020)
- ORI (1994), "ORI policy on plagiarism", *ORI Newsletter*, Vol. 3 No. 1, available at: <http://ori.dhhs.gov/policies/plagiarism.shtml>

---

**Appendix. Letter to the editor about copyright in the networked world**

To the editor:

I read with interest Michael Seadle's article in the most recent issue of *Library Hi Tech*, "Copyright in the networked world: gray copyright" (Vol. 26, No. 2, 2008). The article's central point about "shades of risk" that copiers and rights-holders are prepared to take appears to be a useful contribution, and it could well be the subject of an interesting and productive exchange of views.

But in his article's conclusion, Mr Seadle makes a statement that is erroneous, one that is explicitly contradicted by the US copyright statute. He writes: "... but even the common practice of photocopying articles for reserves is infringement in the same sense that every fair use is an infringement – just one with an established defence that is generally expected to stand up in court" (p. 332).

Not so. The first sentence of Section 107 of the Copyright Law of the United States, the "fair use" section, reads as follows: "Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright."

Appearing in a respected journal in the field of librarianship, Mr Seadle's characterisation of fair use might have the unfortunate consequence of discouraging copying that is entirely legitimate under the law. Librarians and others seeking to make copies and use recorded knowledge responsibly should be encouraged to take, not discouraged from taking, advantage of the fair-use provisions of the copyright law, according to the well-known four factors specified in Section 107: 1) "the purpose and character of the use"; 2) "the nature of the copyrighted work"; 3) "the amount and substantiality of the portion used"; and 4) "the effect of the use upon" the copyrighted work's "potential market" (see "Copyright Law of the United States of America and Related Laws Contained in Title 17 of the United States Code," Section #107, <http://www.copyright.gov/title17/92chap1.html#107>, accessed 13 August 2008; or "Title 17 – Copyrights", United States Code, 2000 Edition, [Washington, D. C., Government Printing Office, 2001], 628-629).

Thank you for the opportunity to comment.

Sincerely,

Richard Hume Werking, PhD  
Librarian and Historian  
Annapolis, MD

**Editor's response**

While I agree with Mr Werking in not wanting to discourage librarians and others from exercising their fair use rights – far from it – he trusts too heavily in the language of the preamble to section 107 (the fair use statute). The courts have shown repeatedly that infringement judgments depend on the four factors in 17 USC 107 and that the preamble plays little role.

Copying for reserves, though a well-established practice, is unfortunately potentially vulnerable because of its effect on the market for a work. Librarians should remember that none of the fair use rights and scholarly exemptions in other countries are permanently secure and need active defence through national librarian organisations.

**About the author**

Michael Seadle is editor of *Library Hi Tech*. He is also a professor at Humboldt University in Berlin, Germany, and director of the Institute for Library and Information Science. He is not a lawyer, and nothing in this column should be considered legal advice.